DOCUMENT RESUME

ED 267 976 SE 046 498

AUTHOR Brenner, Mary E.
TITLE Arithmetic Achievement at Ka Na'i Pono, 1984. Results from Standardized Testing. Technical Report No. 126.
INSTITUTION Kamehameha Schools/Bernice Pauahi Bishop Estate, Honolulu, HI.
PUB DATE Aug 85
NOTE 40p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Computation; Educational Research; Elementary Education; *Elementary School Mathematics; Geometry; *Mathematics Achievement; *Mathematics Instruction; Measurement; Primary Education; Problem Solving; *Reading Skills
IDENTIFIERS *Hawaii; *Mathematics Education Research

ABSTRACT

Hawaiian and part-Hawaiian children in kindergarten through grade 3 were administered the Metropolitan Mathematics Instructional Test to determine their achievement level in mathematics and areas of particular strength and weakness. Only the first grade exceeded the national norms. However, the kindergarten had covered only 21% of the test topics, while the other grades had covered between 54% and 65%. For all classes, computation and geometry/measurement were the areas of greatest strength, while the concept parts of the test were the most difficult. Boys and girls had equal levels of achievement, except that boys did better on word problems in grades 1 through 3. Lower-income children had lower achievement in general, but the gap narrowed in the higher grades; the lower-income groups had superior problem-solving skills in grades 2 and 3. Reading and mathematics achievement were highly correlated, with reading most highly correlated to the non-computation parts of the test. An item analysis indicated that the children had computation, time, and money skills that exceeded their classroom instruction. (MNS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED267976

SE

SE 046 498

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Technical Report No. 126

ARITHMETIC ACHIEVEMENT AT KA HA'I PONO, 1984
RESULTS FROM STANDARDIZED TESTING

Mary E. Brenner

August 1985

Center for Development of Early Education

Kamehameha Schools/Bernice Pauahi Bishop Estate
1850 Makuakane Street, Honolulu, Hawaii, 96817

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Sheida White

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Technical Report No. 126

ARITHMETIC ACHIEVEMENT AT KA HA'I POHO, 1984
RESULTS FROM STANDARDIZED TESTING

Mary E. Brenner

August 1985

Published by the Center for Development of Early Education, 1850 Makuakane Street, Honolulu, Hawaii 96817. The opinions expressed herein do not necessarily reflect the position, policy, or endorsement of the Kamehameha Schools/Bernice Pauahi Bishop Estate, or of the Editor.

All rights reserved. No parts of this report may be reproduced in any form or by any means without the prior written permission of the Kamehameha Schools/Bernice Pauahi Bishop Estate.

TECHNICAL REPORTS
OF THE CENTER FOR DEVELOPMENT OF EARLY EDUCATION

Editor

Sheida White

Editorial Policy Advisory Board, 1984-85

Claire Asam
Rod Calkins
Dale Farran
Debbie Harris
Cathie Jordan
Linda Kobayashi
Sheida White
Tom White

In addition to members of the Editorial Policy Advisory Board, the Editor wishes to thank Burt Furuta and Paula Levin for contributing their time and thoughtful evaluations.

Technical/Production Editor

Joyce Metzger

Darlene Plunkett and Carolyn Lim assisted in the preparation of this manuscript. Their contribution is gratefully acknowledged.

The Center for Development of Early Education is a research and development program of the Kamehameha Schools/Bernice Pauahi Bishop Estate. The mission of the Center is the development of effective methods for improving the education of Hawaiian and part-Hawaiian children, from the prenatal period through the elementary school years. This task is carried out by an interdisciplinary group of educators, psychologists, anthropologists, and linguists, engaged in curriculum development, teacher training, program dissemination, and evaluation. A list of papers, journal reprints, and videotapes is available from Joyce Metzger, Center for Development of Early Education, Kamehameha Schools, 1850 Makuakane Street, Honolulu, HI 96817.

Arithmetic Achievement at Ka Na'i Pono, 1984

Results from Standardized Testing

The Kamehameha Elementary Education Program (KEEP) has developed a reading program that effectively teaches Hawaiian children how to read. Now research attention is being given to the content areas, including math, science and writing. This report is the second in a series exploring the mathematical achievement of KEEP children. The first report (Brenner, 1984) analyzed the standardized test scores in mathematics of KEEP children during the mid-seventies. Before the KEEP reading program was developed, the children performed below national norms in mathematics with average class performance ranging between the 12th and 42nd percentiles. Boys and low income children had particularly low test results. After the KEEP reading program was put into effect, the class averages on the standardized test ranged between the 61st and 93rd percentiles, a clear improvement over earlier testing performance. In addition, the boys and low income children performed at a level much closer to their classmates. However, the children in both the KEEP and non-KEEP reading programs showed similar strengths and weaknesses on the sub-sections of the test. Computation was a strong point in all classes while the children had difficulty with the concept sections. Word problems were of intermediate difficulty.

This report seeks to update the earlier test results and to extend the analyses with children currently enrolled at the KEEP laboratory school, Ka Na'i Pono. The data from this test administration will act as a baseline for further research. Beside presenting global comparisons to the national norms, specific topics will be examined to find the children's relative strengths and weaknesses. Reference will also be made as to which topics had been covered in class. Several analyses are presented to reveal the role of reading skills in math achievement.

Methods

Description of Test

The Mathematics Instructional Test in the Metropolitan Achievement Test (MAT) series 1978, was administered to all classes. This test was chosen because (a) it has questions and scales for the range of topics that would be useful for later research efforts, (b) it has adequate national norms for each grade, (c) it has tests for the entire range of grades at KEEP (K-3), (d) the children have experience with the Metropolitan Reading Test and thus the test format would be familiar, (e) each topic on the test had more than one question, thus allowing for more reliable analysis.

The MAT is divided into five major sections at the elementary level. The Numeration section tests children on conceptual skills including counting, more/less, place value, and number lines. The Geometry and Measurement section includes questions on geometric shapes, money, time, general measurement concepts and the calendar. The Problem-Solving section contains word problems of varying complexity. The Operations section presents computation problems. The Operations: Laws and Properties section tests children on the abstract concepts underlying computations including the role of zero, inverse operations and relational symbol meanings.

All classes were tested on the first four sections. However, in kindergarten, the Metropolitan Instructional Test combines Operations and Problem-Solving into one section and provides national norms for that total score. Only the third grade test included the section on Operations: Laws and Properties. For this reason, that section will not receive much attention in the following analysis.

The role of reading varies across the tests. All grade levels are required to read by themselves the computation problems. In kindergarten, all other sections are read aloud to the children. In first grade the children read six of the word problems for themselves and the other twenty-one problems were dictated by the tester. By third grade, the children read all sections of the test for themselves except for the first 12 of the 30 word problems which were dictated to them by the tester.

The questions of the MAT are organized into objectives such as addition with sums to ten, subtraction of two digit numbers when regrouping is required and solution of word problems when extraneous information is given. Each objective is tested by three multiple choice questions. A child is considered to have competency on an objective when he/she is able to answer two of the three questions correctly. Thus skill on specific topics must be considered in terms of the 2-question criterion for each objective. However, scores on sections such as Operations or Problem-Solving are the sum of all questions answered correctly without regard for performance on specific objectives.

The objectives overlap on different sections of the test. Thus both Operations and Problem-Solving have questions covering addition to ten in first grade. The numerical content of questions can be held constant, thereby aiding analysis on how problem structure affects performance.

Description of Sample

Form J of the MAT was administered to all classes (K-3) at Ka Na'i Pono School in the spring of 1984. These classes are not strictly comparable to previous KEEP classes nor to each other due to variations in sampling procedures. In the mid-seventies KEEP students were more "at-risk". They tended to be lower-income and were recruited in large part from the

rolls of the Department of Social Services and Housing. The current students were chosen from the larger pool of children who had applied for admission to Kamehameha Schools. While many are from lower income families, they are considered less "at-risk" than earlier groups. Individual classes also have distinctive characteristics. The third grade was recruited mostly from rural Oahu. Because of these sampling variations, comparisons between classes must be considered tentative. In addition, the generality of these findings to any larger population of Hawaiian children must be considered limited at this point.

Test Administration

The MAT was administered to each class in several sessions over the period of a week. The researcher did all test administration while the classroom teachers served as proctors. All classes were given practice problems before the actual test sessions to acquaint them with the test procedures. The children were given slices of apple after each test session. First graders were given apple slices before the test as well, since they were tested in the morning and it was not clear that all students had eaten breakfast.

Makeup tests were given to children who had missed individual sections of the test but not those who missed the entire test. As a result the completion rate for each class was very high. Table 1 summarizes the number of students in each class, classified by whether or not they took the entire test.

Table 1

Number of students by grade level and test completion status

Grade	No Test	Incomplete Test	Complete Test
K	0	2	28
1	2	0	29
2	0	0	26
3	0	0	29

Factors in Test Achievement

A class' achievement on a standardized test is the outcome of a variety of factors. While all these factors will not be systematically analyzed here, some information will be given to put the test results into perspective, and to provide suggestions for future research.

Curriculum. The Metropolitan Achievement Test is fairly representative of the topics covered on any standardized test. However, there was not an equal degree of correspondence between the curriculum and the test items for the classes tested. All classes used the Addison-Wesley math textbooks, 1981 edition, and these texts served to structure the curriculum at each grade level. However, each class differed in how far through the text they progressed and in what supplementary topics were covered. To evaluate the degree of correspondence between the classwork and test items, each teacher was asked to evaluate each objective in the math test given to her class. A topic was rated as "covered" if the entire class had been instructed in it in a form comparable to the test questions. A rating of "some covered" was given if only some students had covered the topic or the test presentation of the concept differed substantially from

the class presentation. A topic was considered "not covered" if it had not been a part of lessons in math or any other content area.

Table 2

Correspondence between Test Items and Curriculum

Grade	Numeration	Geometry	Problem-Solving	Operations	Total
	(%)	(%)	(%)	(%)	(%)
Kindergarten					
Not Covered	50	33	100	100	57
Some Covered	17	44	0	0	21
Covered	33	22	0	0	21
First					
Not covered	20	22	0	22	16
Some covered	30	0	44	11	22
Covered	50	78	56	67	62
Second					
Not covered	40	50	20	9	29
Some covered	0	0	40	27	17
Covered	60	50	40	64	54
Third					
Not covered	20	27	20	0	17
Some covered	10	13	40	18	19
Covered	70	60	40	82	65

Note: Not all columns sum to 100% because of rounding.

Figure 2 shows that the kindergarten class experienced the worst fit between the test and curriculum. Quantitative topics were not addressed at all in the kindergarten class, because the text emphasized the concepts underlying counting and classification. Several other topics were covered only in science or math lab and not in the regular math period. Overall, only 21% of the text objectives were covered in class. The other grades were more comparable in their coverage of test objectives, with totals ranging between 54% and 65%. Across the 3 grades, operations tended to be the best covered type of problem while problem-solving was relatively de-emphasized. There were specific test objectives that were not covered very much in any class. These included expanded numbers, calendar skills, missing addend/subtrahend computations, number sentences and more advanced word problems. Overall, the test was considered an adequate assessment of classroom progress for all classes except kindergarten. The kindergarten test scores must be interpreted with caution.

Test-taking Situation. The actual testing situation itself can be a factor in how well the test scores represent student levels of achievement. The children in this study had experienced standardized testing before, so the testing procedures were familiar to them. However, there were some problems apparent in these testing sessions.

The tests were administered during the last two weeks of school. While this meant that the KEEP children had more of the school year to cover the test topics than the standardization (national norm) comparison groups, the testing situation itself was problematic. The children were excited and agitated because the end of the school year was at hand. Instruction had been stopped in some subjects and had been replaced by more recreational activities. Thus the children may not have been as

attentive to the test as they could have been had they been tested earlier in the year.

All classes, except first grade, were tested in the afternoon during their normal math period. As a result, the children were somewhat tired and a few even fell asleep during the test. Also, discipline problems tended to be greater in the afternoon. In the future, math testing should be done in the morning. The children's concentration will be better and results will be more comparable to the reading test.

The test procedures themselves were not a major problem in any class except kindergarten. Except for a few students, the time limits were more than adequate. The older students had little difficulty in following the testing procedures. However, students in the kindergarten class were frequently confused during the test. They were the only students who had never taken a Metropolitan test in any subject. The test administrator and proctor had to monitor each child before every question to make sure they were on the appropriate question and that they were marking their answers correctly. Combined with the difficulty of the test, the testing sessions were quite trying for some of the kindergarten children. Special care should be used in the choice of a test for the kindergarten class in the future. However, their testing session provided some interesting insight into kindergarten math achievement, and standardized testing should not be abandoned altogether at this level.

In order to obtain a general idea of the test's validity, the teachers were asked to examine the test results in terms of the students' class performance. Overall the teachers felt that the test represented the relative achievement of specific students. A few students did worse than expected in part because they were slow test takers.

Teaching Style. It is impossible at this point to relate specific teaching methods to achievement on the math test. Only superficial observations have been done to date of how math is taught at KEEP. The classes vary in many ways including the degree to which a center approach has been adopted, how many manipulatives are used and how much time is spent on math. The kindergarten and third grade classes had long-term substitute teachers so teaching methods were not constant throughout the year. A math lab was instituted at mid-year with the goal of incorporating more manipulative materials into instruction. This goal was only partially realized because of time constraints. The math lab did serve a valuable function in terms of introducing some topics such as money and thereby helped the students with the MAT.

Other Skills. Achievement on a standardized mathematics test relies on more than what a child learns in math class. Reading skill becomes a major part of the MAT by third grade. The role of reading in math achievement will be addressed in a later section.

Science also teaches math skills at the elementary level. A number of topics, particularly measurement, were presented more often in science lessons than in math. In the future, a coordinated effort in science and mathematics may be quite productive.

Children also bring relevant skills from non-school settings. Some of the test objectives such as telling time and money usage clearly relate to activities done outside of school. Others such as computation are typically learned in school but some children are taught these at home as well. Specific test objectives were examined to see if the children have strengths in non-school skills. These results are presented later in this paper.

Results

General Test Results

There is much variation in the overall math achievement of the four classes. Figure 1 show the percentile for the average performance in each class. The first grade is the only class that exceeds the national average percentile of 50%. The very low average performance in the kindergarten class is readily explained by the lack of match between the class curriculum and test objectives. Otherwise, the test results do not vary in accord with how well the test matched the topics covered in class. These results are much lower than those obtained among KEEP classes in the 1970's, but in both cases, first grade did markedly better than the other grade levels.

Figure 2 shows how each class performed on each section of the test. In every grade except kindergarten, the children's greatest strength is in operations. This follows from the greater number of computation objectives that were met in each class. The textbooks also provided a lot of practice in computation.

Children had the lowest level of achievement in Numeration, although numeration objectives were fairly well covered in each class. The classes tested in the 70's also had the most difficulty with this section of the test as well.

The results for the other 2 sections of the test are mixed. Both the first graders and third graders were fairly competent in Problem-solving--both averaged close to the national average. The second graders found this section quite difficult and averaged at only the 29th percentile.

The first grade was quite strong in Geometry and Measurement. The first graders had covered 78% of the test objectives in class and this

probably contributed to their good performance. The Geometry and Measurement section was the kindergarten class' strongest part of the test. Although they averaged at only the 28th percentile, they accurately answered 66% of the questions. Since only 22% of the topics on this section had been covered in class, it appears that the Kindergarten children were using skills that they had learned in other settings.

Figures 3 and 4 compare the performance of boys and girls. In every grade except third, they had a identical overall performance. In third grade the boys had a slight advantage that could be due to chance. However, the testing in the mid-1970s also found a slightly better overall performance among boys who had had the KEEP reading program in second and third grade. On the sub-sections of the test, there was no clear pattern of results by sex except in problem-solving. In every grade except kindergarten, the boys did better in problem-solving. This was also found in the earlier KEEP classes as well as in many other studies (Fennema, 1981: Marshall, 1984), although it is not generally true for children in Hawaii (Brandon, Newton & Hammond, 1985). Overall stronger performance by boys in problem solving occurred even though boys tend to have lower reading skills.

A comparison by social class is included because one goal of the KEEP research program is to design educational progams for "at-risk" children. The measure of social class used here is whether the child's family receives an income from the Department of Social Services and Housing (DSSH). This is not the optimal measure of social class because some of the other children are from low-income families. However, this measure was readily available and quite straightforward in its application. DSSH children performed at a lower level in every grade but the gap narrows as

grade level increases (Figure 5). The KEEP reading program may help to narrow the very large early differences between DSSH and non-DSSH children represented by the 26 point percentile difference in kindergarten. In Figure 6, it can be seen that the DSSH children performed lower on every section of the test except Problem-solving. In both second and third grade, the DSSH children exceeded their non-DSSH peers in problem-solving. It may well be that the KEEP reading program begins to pay dividends for these "at-risk" children when reading begins to be applied in math. The DSSH children also did better than their peers in third grade on the Operations: Laws & Properties section of the test.

Arithmetic Concepts

The Numeration section of the Metropolitan Instructional Test was designed to evaluate children's conceptual understanding of arithmetic. Although the topics on this section were fairly thoroughly covered in each class, this section was typically the most difficult for the Ka Na'i Pono children. Some topics, such as expanded notation, were difficult for all grades because they were not covered in class, and other topics caused difficulty despite class coverage.

The most striking finding is that the kindergarten and first grade children had extreme difficulty with questions involving the phrases "more than" and "less than". The kindergarten class missed 70% of the questions on these terms while the first grade class missed 57%. These figures are well below the national error rates of 36% and 30% on the same questions. The second grade class was closer to the national norm but still slightly below it. The kindergarteners did not display any coherent pattern of errors in their wrong answers. However, the first grade class tended to interchange the meanings of "more than" and "less than".

The other concept that was exceptionally difficult was one which required the children to find a number between two others in value. The first grade class was 14 percentile points below the average national percentile and the second grade was 27 points below the national average. Typically the children chose a number that was one larger than the larger of the two establishing a range. In other words, the children tended to see the question as a counting task.

The causes for these particular difficulties are not readily discernible from a multiple choice format test. The wording of the questions themselves may have caused confusion, or the local form of English may be interfering with the standard English forms of these phrases. Or there could be underlying difficulties in numerical concepts. However, the third grade class had no difficulty with these questions, indicating that the lower grade results represent lagging performance and not necessarily a major deficit.

All of the classes showed strength in basic counting skills and number names. Performance was inconsistent across grades on numerical terminology including ordinal number names, odd/even numbers and quantitative adjectives. Since performance on the Numeration section was correlated with reading skill, it is likely that the children's performance on this section was partly a function of their general language skills, including listening comprehension and vocabulary development.

Specific Strengths and Weaknesses: Item Analysis

Each section of the Metropolitan Achievement Test is made up of objectives which specify the component skills being tested. While the amount of detail can be overwhelming at this level of analysis, there is important information that can be gleaned only by examining the component

skills. As the basis for further research, an analysis was done to find out how well the children did on skills which are used in non-school settings. These skills are found primarily in the Geometry and Measurement section, and secondarily on the Operations section. The match betweeen curriculum and test questions can also be more clearly explicated through this analysis.

To examine component skills, an item analysis was done of each section of the test. In an item analysis, the number of errors for each separate question is summed for the entire class. The percentage of the class missing a question is then compared to the percentage of the national sample missing that question. The class performance for each objective is the mean percentage of errors for the three questions which test the objective. The class' relative proficiency on each concept is then evaluated in two ways. The skills were rank ordered in terms of how many children missed each question, giving a picture of the children's relative mastery of each skill. The skills were also rank ordered in terms of how much better or worse than the national norms the class did on each objective. This controls for the varying levels of difficulty across skills. As an example, the second graders may have missed more multiplication computations than any other. But since this is an advanced skill for second grade, their greater number of errors in multiplication can not really be considered a major problem unless other second graders throughout the nation had a much greater mastery of multiplication.

Kindergarten Item Analysis. The kindergarten class represents the clearest instance of how non-school learning experiences influence skill development. Because so few of the test objectives were taught in the kindergarten curriculum, their competencies must have derived from home,

the preschool or other experiences. The Problem-Solving and Operations section provides the clearest evidence because none of the skills on this section were taught in the kindergarten class. Table 3 shows what percentage of the class met the competence criterion (2 of 3 questions correct) for each skill.

Table 3

Percent of Kindergarten Class showing Competence in Computation

Skill	% Reaching Criterion
Adding and Subtracting 1, word problem	72%
Addition to 18, computation	41%
Addition to 10, computation	28%
Addition to 18, word problem	24%
Subtraction to 10, computation	21%
Subtraction to 10, word problem	17%
Addition to 10, word problem	10%

Most of the children had some comprehension of basic operations as shown by the fact that 72% could solve the simplest word problems. Perhaps more surprising is the relatively high percentage that could solve simple computation problems. Forty-one percent were able to do addition problems with sums to 18. This involves facility with the symbolic representation of the addition operation and the ability to read what are fairly large numbers for this age group. Two other observations can be made from Table 3. First, the computation problems are relatively easier than the word problems. For instance, 41% of the children could do addition computations to 18 while only 24% could do word problems with sums to

18. Secondly, the addition problems with larger sums were answered by more children than the problems with smaller sums (to 10). This was true for both computations and word problems.

Despite their surprisingly strong calculational skills, these children did not meet the national standards for any objective on this section of the test. They were only 6% below the national norm for sums to 18, but 54% below the national norms for word problems with sums to 10.

Although the kindergarten class was well below the national standards in Geometry and Measurement as a whole (mean percentile = 28%), they had some strong skills as well. As a class they exceeded the national norms on telling time and equalled them on both money objectives. Neither of these topics were covered in the text, although the children received some instruction in math lab on money concepts and some supplementary exercises with clocks. The kindergarten children had the most difficulty with questions regarding measurement tools (except for clocks) and fractions.

The overall impression from the kindergarten class is that the children have substantive quantitative skills that have not been adequately stimulated by the school curriculum. Their skill with money and clocks indicates the practical basis of at least some of their skills.

First Grade Item Analysis. The first grade class performed above the national norms on every objective in the Operations section of the test. Compared to the national norms, their strong points were subtraction of two digit numbers, problems with three addends, missing addend problems and basic subtraction. These are all the more advanced skills in this section. Some of these topics, such as problems with missing addends, were not even covered in class, indicating that the children's solid understanding of addition was enough to give them some insight into a

different type of problem. In absolute terms, these same difficult items were the hardest for the children. Only about half of the class was able to do the kinds of problems listed above. The children's performance on word problems was similar to that in Operations. They were above the national norm for every objective except problems involving two digit addition and subtraction. Problems involving number sentences gave them some difficulty, in part because they had little experience with number sentences in class.

On the Geometry and Measurement section of the test, the first grade children exceeded the national norms on six of nine objectives. Like the kindergarten class, they were particularly strong in money and time skills. The calendar questions were by far the hardest for the children. They missed 70% of the calendar questions and were 21 points below the national percentiles. However, about 30% of the children had some calendar skills even though the calendar had not been a part of the arithmetic lessons. Other questions that were hard for them involved reading the names of geometric figures and reading the words representing time on the clock. This suggests that the children were not familiar with the written vocabulary of arithmetic class.

Second Grade Item Analysis. The second grade exceeded the national norms on four of the eleven objectives in Operations. They were strong in basic number facts in addition, subtraction and multiplication. They were weakest in subtraction skills, particularly two digit subtraction. Their skill in multiplication exceeds what they were explicitly taught in class. They got 60% of the multiplication questions with numbers up to five correct even though only the top group had covered this topic in class. Forty percent of the questions on more advanced multiplication

questions were answered correctly, even though none of the children had covered the topic in class. The results for Problem-Solving are the inverse of those for Operations. The second graders were closest to the national norms in solving the more difficult problems such as two digit addition and word problems with extraneous information. However, they were somewhat below the national norms for every objective in this section. Although the number sentence format of the test was somewhat unfamiliar, the word problems requiring only a number sentence answer were answered correctly slightly more often than problems requiring the answers. Overall, they got 64% of the word problems correct.

The second graders exceeded the national norms on three of seven objectives in Geometry and Measurement. Like both the kindergarten and first grade, they were particularly strong in basic money and time problems. They were also strong in fractions. However, they were weak in more advanced time and money concepts, particularly those involving written vocabulary. The calendar was their weakest point of all. None of the advanced questions in time, money nor the calendar had been covered at all in class. About 25% to 30% of the questions in these sections were answered correctly, primarily because a number of children in the class met the two question criterion for competency rather than through a pattern of guessing. Half of the class met the criterion for the calendar questions even though calendar was not a topic in the curriculum.

Third Grade Item Analysis. The third grade class exceeded the national norms on all but one objective in the Operations section of the test. They had difficulty with multiplication with regrouping probably because this topic had not been covered in class. This class was strong in multiplication and division facts to nine, as well as adding 2-digit

numbers. As a class, they averaged 70% or better on every objective except division facts between five and nine and multiplication with regrouping.

This class exceeded national norms in eight of ten problem-solving objectives. They had trouble with number sentences representing multiplication and division problems and multiple-step problems. In general, answering with a number sentence proved harder for the children than actually solving a problem. In absolute terms, word problems with multiplication and division were much harder than those for addition and subtraction. These topics had only been partially covered in class. However, the children did a bit better on complex word problems involving several steps or with extraneous information and these were not done in class at all. Thus the children's difficulties with multiplication and division word problems derive from their still incomplete understanding of these operations, including relevant vocabulary, rather than from a more general difficulty with the genre of word problems.

The third grade had more difficulty with Geometry and Measurement, exceeding national norms only on two out of 15 objectives, and about equalling them in four others. They were strong on problems with coins and in calendar questions. They were very weak in fractions and had difficulty with all of the measurement problems. The third grade test in this section is substantially harder than it is in the other grades. Both the children at KEEP and the national standardization groups had difficulty in the measurement areas with less than half of either group being able to answer each question. Some of the questions, particularly the metric measurement questions, were not covered in the third grade class. The third grade was the only class which had to read the Geometry and

Measurement section. Thus reading skill, particularly in the specialized vocabulary, may have been a factor here.

Item Analysis Summary. The pattern across grades for Operations and Problem-Solving is very mixed but much more consistent in Geometry and Measurement. Some of the classes were strongest in basic facts while others were relatively stronger in the more advanced operations. In every class, there were children competent in each Operations objective, whether or not the objective had been covered in class. One particular aspect of the Problem-Solving test format was more difficult for the KEEP children than the national norm groups. Among the norm groups, number sentence questions were routinely easier than questions requiring a calculated answer. However, the KEEP children often found these questions harder. Number sentences were not consistently covered in the texts used at Ka Na'i Pono. It is an empirical question whether number sentences actually help children to learn how to solve word problems. In general the Addison-Wesley texts used in these classrooms de-emphasize word problems. However, questions with number sentences formed about a quarter of the word problem questions on the Metropolitan test.

In all classes, the children were strong in basic money and time questions. By third grade, a number of children had also acquired basic calendar skills. The consistency of these results across grades, despite varying amounts of emphasis in the classroom, probably represents a common emphasis on these topics in non-school settings. Whether the children have actually acquired these skills outside of school or merely have a greater interest and receptivity because of their potential utility is an open question.

The children also have difficulty in measurement in kindergarten and third grade. Measurement is not covered in first or second grade test so it's not clear if this is a general weakness across the Ka Na'i Pono classrooms.

The KEEP children also had some difficulty in questions requiring them to read specialized arithmetic vocabulary. While they could point out the shape when given the name, it was harder for them to pick the name when shown the shape. They had the same difficulty with money problems. They could identify coins and read their numerical values. But the questions involving the names of coins--"quarters", "pennies", etc.--proved harder.

The children's problems with mathematical vocabulary and their superior facility in computations as compared to other sections of the test raises questions about the importance of reading in arithmetic. Given that the children had to read some sections of the test by themselves, difficulties with reading may have obscured some math competencies.

Reading Skills and Arithmetic Achievement

The role of reading achievement in math achievement is explored in two ways. First, scores on each section of the math test are correlated with a recent measure of reading comprehension (Table 4). The second method involves comparing problems of similar numerical difficulty across modes of presentation--pure numerical, listening and reading (Table 5).

Table 4

Correlations Between Math Achievement and Reading Comprehension

Grade	Correlations				Variance Accounted for in Math Scores			
	Numeration	Geometry Measurement	Problem Solving	Operations	Numeration	Geometry Measurement	Problem Solving	Operations
1st	.58	.74	.59	.56	34%	55%	35%	31%
2nd	.75	.74	.70	.53	56%	55%	49%	28%
3rd	.69	.73	.65	.53	48%	53%	42%	28%

For all classes, the correlation between reading and operations is the lowest while that between reading and Geometry and Measurement is the highest. The correlation between the other sections (Numeration and Problem-Solving) and reading are higher than those between reading and Operations, although the differences in first grade are not great. The right half of the table highlights the relationship between reading and math in a slightly different way. These percentages show how much of the variation between scores in math can be accounted for (statistically) by variation in reading scores. For example, 31% of the variation in computation skills (Operations) in first grade can be accounted for by children's differences in reading skills while 55% of the variation in Geometry and Measurement is accounted for this way. In other words, the Geometry and Measurement section shows almost twice as much relation to reading as does Operations.

All of the correlation coefficients in this table are statistically significant. However, skill in reading per se may not be as strong an

explanatory factor as these figures may seem to indicate. Both reading and math achievement are dependent upon children's basic ability and their degree of engagement with school (Aiken, 1972). The correlations between Operations and reading comprehension can be taken to roughly represent this underlying commonality. If so, the reading scores then explain only about another 20% to 30% of the variance in math. Another factor that must be taken into consideration is that only the third grade was required to do substantial amounts of reading on any section of the test besides problem-solving. Even taking these factors into account, all grades show an important relationship between reading and math on other sections of the test.

A comparison of specific questions in the Problem-Solving section and the Operation section demonstrates this relationship further. On the Problem-Solving section of the test, every class was read some of the questions but had to read some others for themselves. Comparing of these modes of presentation reveals how important reading is to problem solution. A comparison to computation reveals the relative importance of general verbal skills as opposed to just skill in reading. Tabel 5 lays out the comparisons for each grade. In general the children made slightly more mistakes when they had to read a problem for themselves. However, these differences are small, about 3% to 6%. This is equivalent to 1 or 2 children in each class having more difficulty answering a question when required to read it alone. A much larger gap exists between the children's skill at solving a word problem and their ability to solve the numerical computations. These differences range from 5% to 20%. However, such differences are found with the Metropolitan national sample as well as in other national samples of children (Carpenter et al. 1980). Thus,

the KEEP children are not generally deficient in their problem-solving abilities. However, they compare less favorably to national samples in problem-solving than in computation so improvement could be made in their problem-solving skills. Since this lag is present in both listening to and reading word problems, their difficulty seems to lie more in developing and practicing strategies for attacking word problems rather than in their reading or computational skills.

Reading skill per se may have some relationship to the high correlation between Geometry and Measurement score and Reading Comprehension. As mentioned earlier, the children had more difficulty with questions which involved choosing words for an answer rather than numbers or a picture. However, these few questions do not explain the rather substantial correlations between reading and geometry. Since only third graders had to read the questions, it seems likely that the key factors in the correlations are actually general listening skills and/or vocabulary skills.

A number of explanations exist for the observed relationship between reading and math achievement for Ka Na'i Pono children. The workbooks in the Addison-Wesley series may have required reading or vocabulary skills beyond the range of some students, thus precluding effective use of the materials. The less verbal children may have been less interested in topics that were presented verbally in class. Computations and everyday skills such as using money and telling time are active and practical skills as opposed to most of the activities in the student workbooks. More use of manipulatives and active projects might involve more students. More of these possibilities can be tested with existing data; each opens a new line of future research.

Table 5

A Comparison of Problems by Mode of Presentation

	Mode of presentation	Problem Content	% Wrong
First	Listen-word problem	Add & Subtract to 18	26.5
	Read-word problem	Add & Subtract to 18	33.3
	Compute only	Add & Subtract to 18	28.5
	Listen-word problem	Number Sentence to 18	24.0
	Read-word problem	Number Sentence to 18	35.0
Second	Listen-word problem	Add & Subtract to 18	29.7
	Read-word problem	Add & Subtract to 18	32.0
	Compute only	Add & Subtract to 18	12.2
	Listen-word problem	Add & Subtract-2 digit	39.7
	Read-word problem	Add & Subtract to 18	39.7
	Compute only	Add & Subtract to 18	33.3
	Listen-word problem	Number Sentence-Sums to 18	24.3
	Read-word problem	Number Sentence-Sums to 18	30.7
	Listen-word problem	Number Sentence-2 digit prob.	24.0
	Read-word problem	Number Sentence-2 digit prob.	29.3

	Mode of presentation	Problem Content	% Wrong
Third	Listen-word problem	Number Sentence-Sums to 18	7.7
	Read-word problem	Number Sentence-Sums to 18	16.7
	Listen-word problem	Multipli. & Division to 9	45.3
	Read-word problem	Multipli. & Division to 9	48.0
	Compute only	Multipli. & Division to 9	37.0
	Listen-word problem	Number Sentences-Multiplication and Division	52.3
	Read-word problem	Number Sentences-Multiplication and Division	44.3
	Listen-word problem	Add & Subtract to 18	10.0
	Read-word problem	Add & Subtract to 18	12.3

Summary and Conclusions

A number of factors influenced the achievement of the Ka Na'i Pono children on the Metropolitan Instructional Test in Mathematics. The most obvious is the degree of fit between the Metropolitan test and the materials children had covered in class. The kindergarten class had covered only 21% of the test topics while the other grades had covered between 54% and 65%. The test taking situation also presented problems. It is suggested that all classes be tested in the morning and earlier in the Spring semester.

The classes varied greatly in their overall level of achievement. Only the first grade exceeded the national norms. For all classes, computation was the greatest strength while the concept parts of the test were most difficult. The overall level of achievement of these classes was

substantially below that of other classes which had been taught with the KEEP reading program in the mid-seventies (Brenner, 1984).

Boys and girls had equal levels of achievement on the test, except that boys did better on word problems in Grades 1 through 3. The lower income children, as measured by whether or not the family received public assistance, had lower test achievement in general but the gap narrowed in the higher grades. The lower-income group had superior problem-solving skills in second and third grades.

Reading and math achievement are highly correlated, in part because both are strongly related to children's underlying abilities and motivation. Reading was most highly correlated to the non-computation parts of the test although the children did not always have to read those sections of the test. Questions presented orally to children were just about as hard as questions the children read for themselves. Thus low scores can not be attributed to reading deficiencies alone. More general language skills and methods of instruction must be considered.

Except for the first grade class, the Ka Na'i Pono children did not meet the Kamehameha Schools' objective of equalling the national average in mathematics achievement. This is in accord with other information about the achievement of Hawaiian and part-Hawaiian children in mathematics--they are below the national norms and below other ethnic groups in the state of Hawaii (Brandon et al. 1984). There is clearly a need for further research to discover instructional methods that will raise the arithmetic achievement of Hawaiian children.

A clear statement about the goals of mathematics instruction is also needed as a guideline for further research. The children in this study were weakest in the conceptual and problem-solving skills that may be most

vital for future achievement in mathematics. Just as the KEEP reading program set the objective of raising reading comprehension, the mathematics program needs direction to guide instructional practice. Since the children displayed some problems with the language skills involved in mathematics, some of the principles which direct the KEEP reading program may be profitably applied to the mathematics program. In particular, ways of relating the children's non-school mathematics experience (which were reflected in parts of the Geometry and Measurement section) may offer the key to raising the children's interest and achievement in mathematics.

References

Aiken, L. R. (1972). Language factors in learning mathematics. Review of Educational Research, 42, 359-385.

Brandon, P., Newton, B. & Hammond, O. (1984). The superiority of girls over boys in mathematics achievement in Hawaii. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Brenner, M. E. (1984). Standardized Arithmetic Testing at KEEP (1975, 1977) (working paper). Honolulu: Kamehameha Schools, Center for Development of Early Education.

Carpenter, T. P.; Corbitt, M. K.; Kepner, H. S., Sr.; Lindquist, M. M. and Reys, R. E. (1980). National assessment: A perspective of students' mastery of basic mathematics skills. In M. M. Lindquist (Ed.) Selected issues in mathematics education (pp. 215-257). Berkeley: McCutchan.

Fennema, E. (1981). The sex factor. In E. Fennema (Ed.) Mathematics education research: Implications for the 80's (pp. 91-110). Alexandria, Va: Association for Supervision and Curriculum Development.

Marshall, S. P. (1984). Sex differences in children's mathematics achievement: Solving computations and story problems. Journal of Educational Psychology, 76, 194-204.

FIGURE 1

Metropolitan Mathematics Test Results

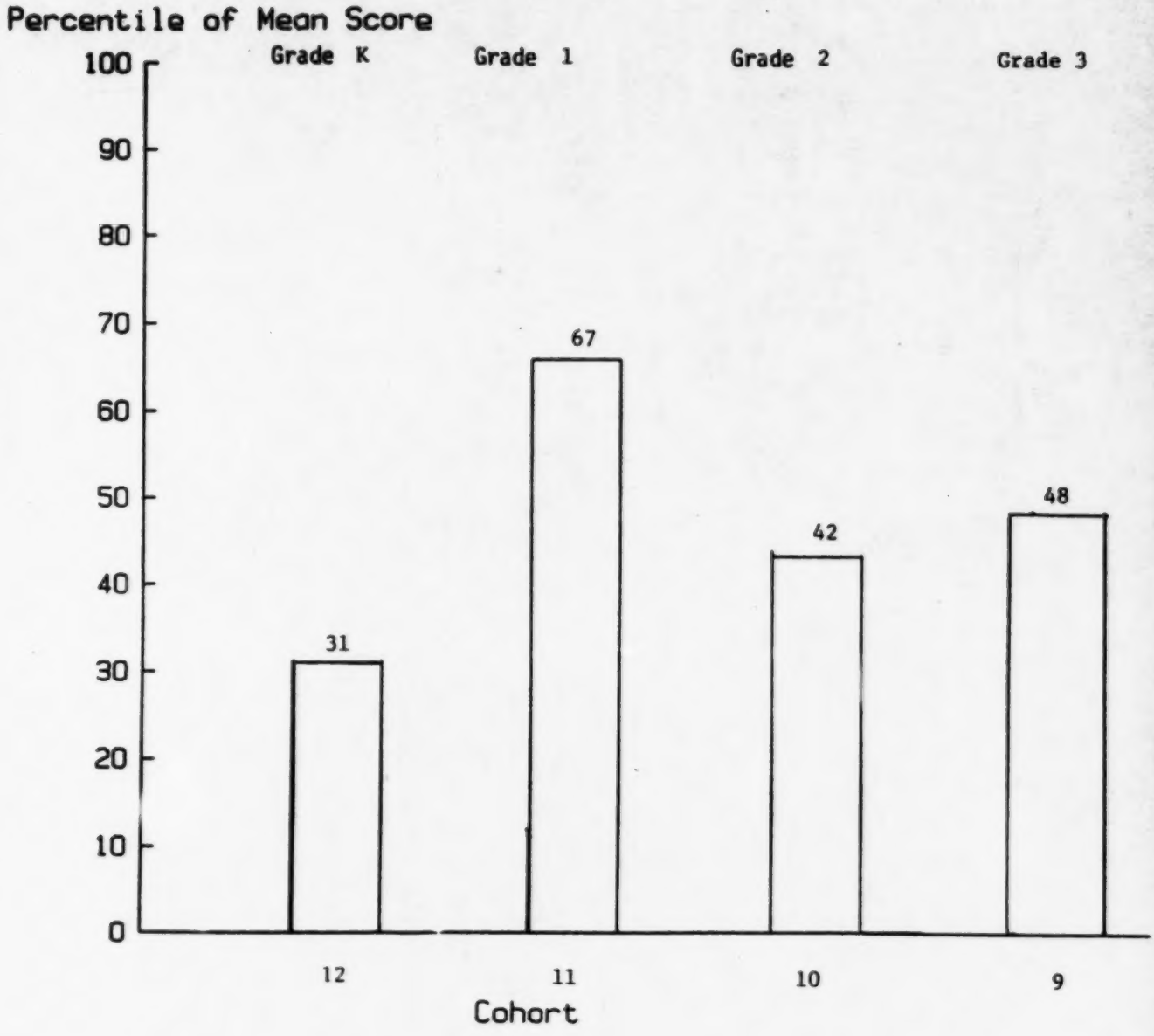

FIGURE 2
Metropolitan Mathematics Test: Comparison of Sub-Sections of Test

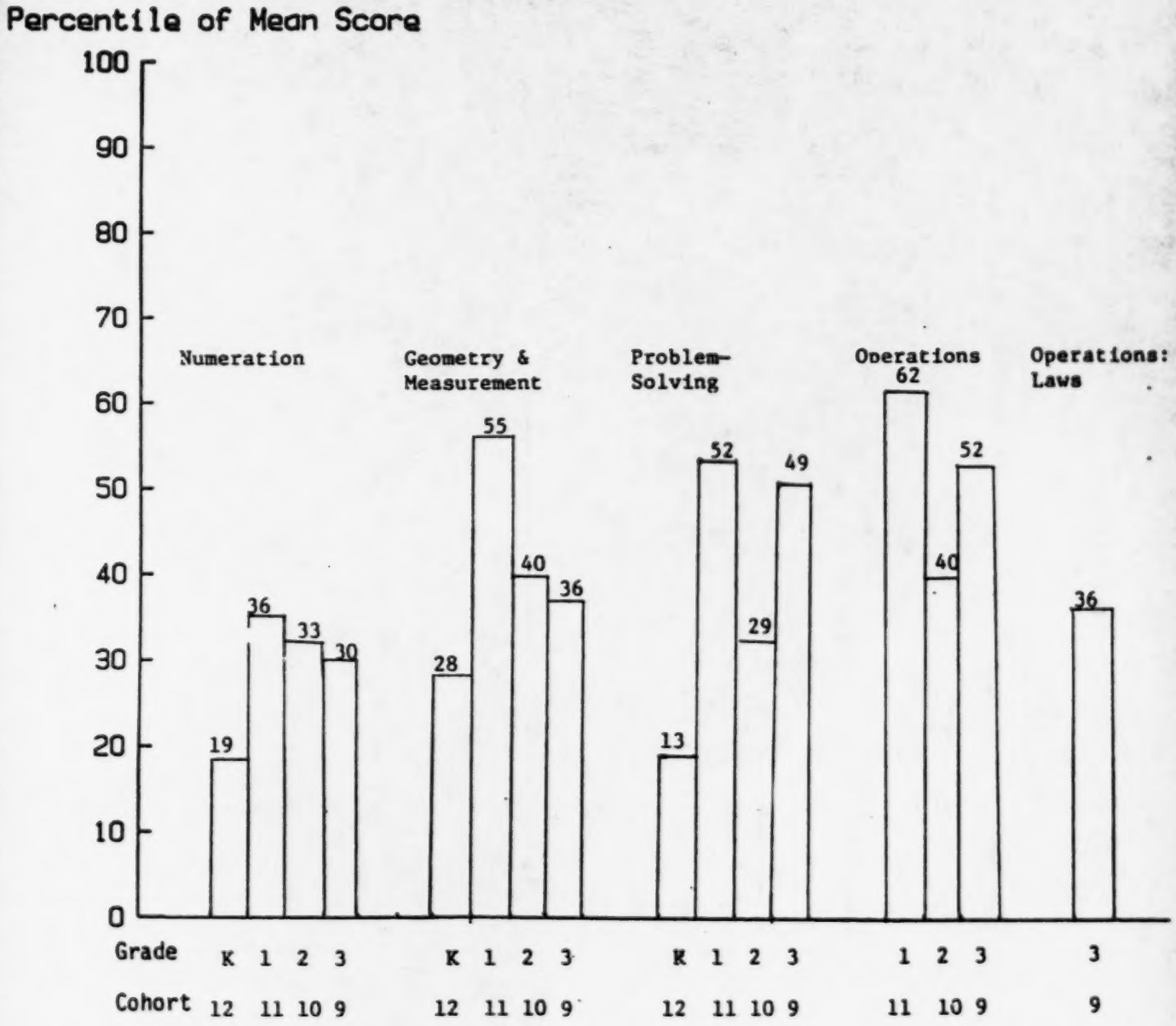

FIGURE 3
A Comparison of Boys and Girls in Arithmetic (Total Test Score)

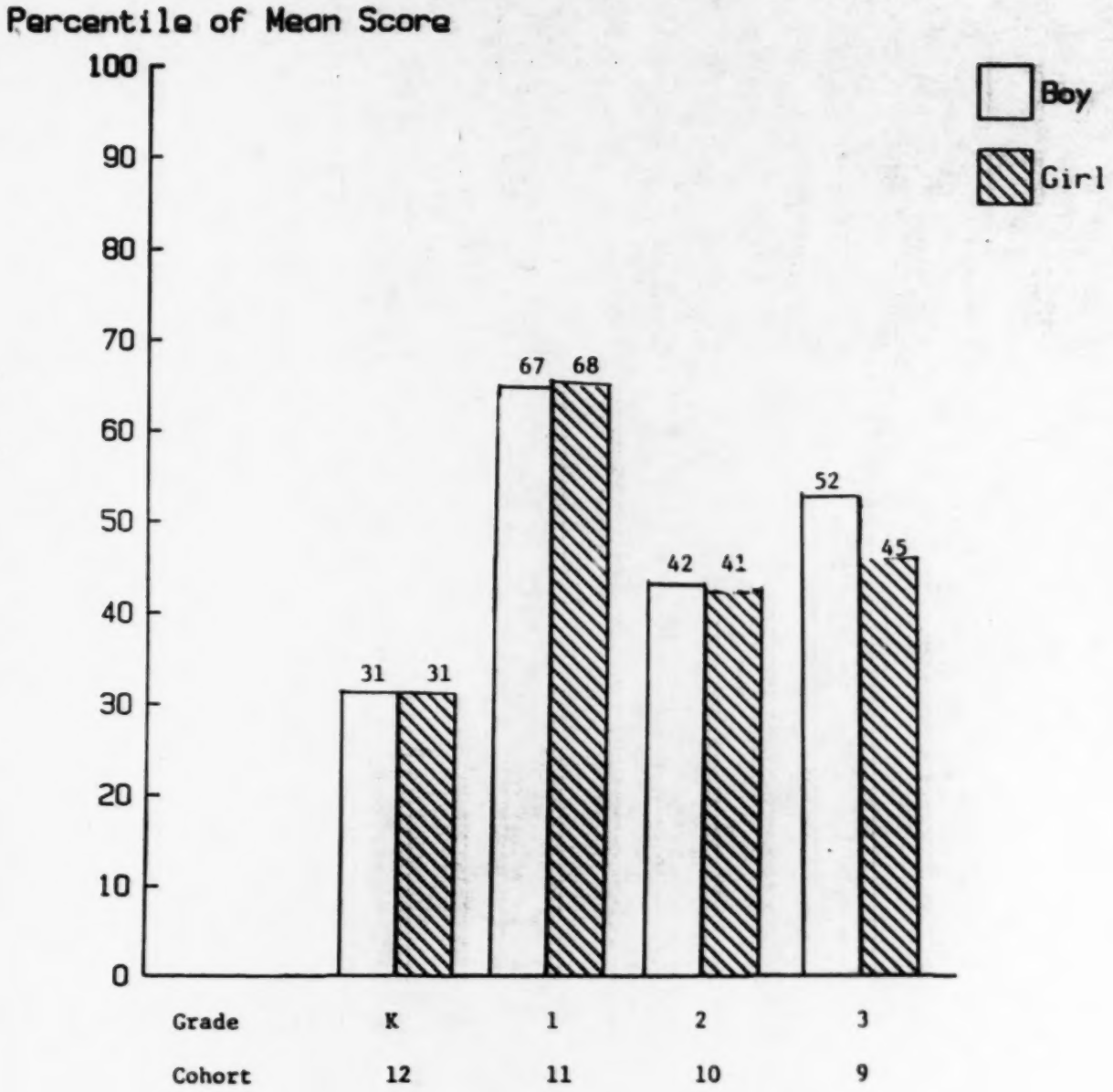

FIGURE 4

A Comparison of Boys and Girls in Arithmetic: Sub-Sections of Test

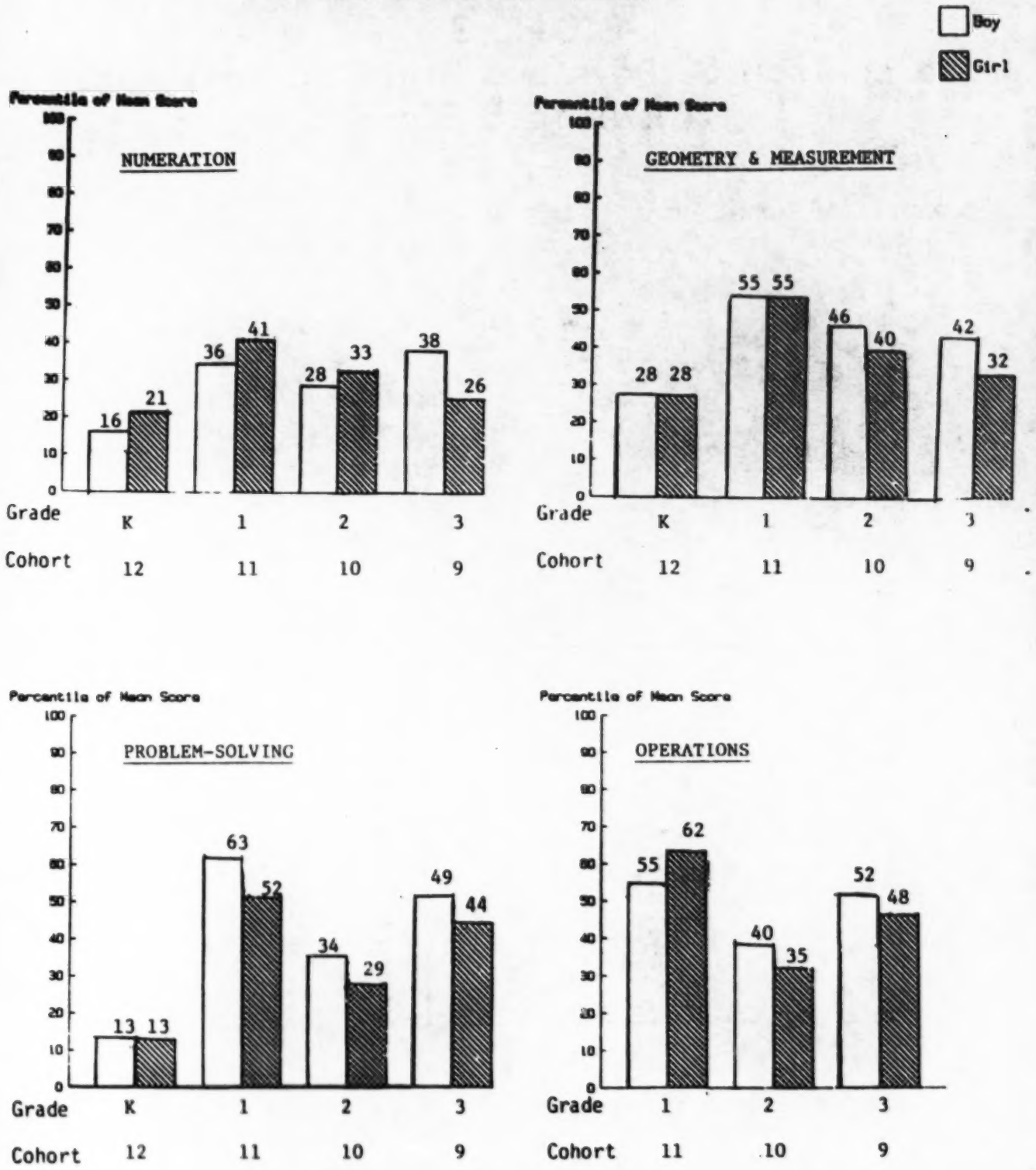

FIGURE 5
Comparison of DSSH Children and non-DSSH Children on Mathematics Test

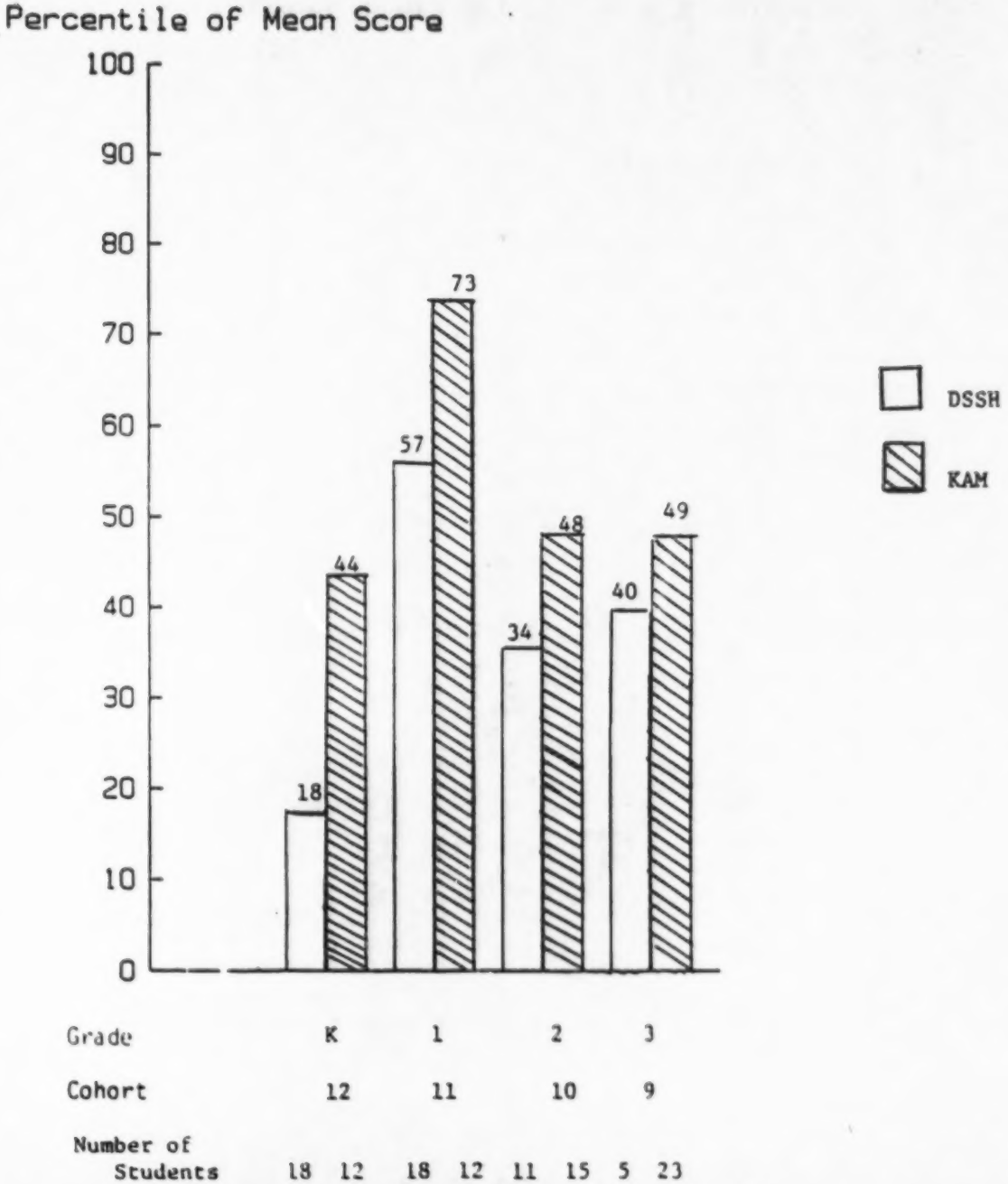

FIGURE 6
Comparison of DSSH Children and non-DSSH Children: Sub-Sections of Test